Computer assisted training in statistics : internet and multimedia : a survey of existing tools

Annie Morin, Gilbert Saporta, Antonis Spinakis

Computer assisted Training in Statistics: Internet and Multimedia: a Survey of Existing Tools.

Annie Morin[1], Gilbert Saporta[2] and Antonis Spinakis[3]

[1] IRISA, Université Rennes 1, F-35042 Rennes Cedex, amorin@irisa.fr
[2] CNAM, 292 rue Saint Martin, F-75141 Paris Cedex 03, saporta@cnam.fr
[3] University of the Aegean, 8 Michalon Str., GR-82100 Chios, aspi@aegean.gr

Summary

During the last ten years, many tools have been produced for self-learning in statistics. It began with distance teaching with paper and pencil, video or television; then specialised software appeared: the first ones on a book model, then using the resources of multimedia. Recent developments are now with Internet. This paper aims to classify available products and analyse future trends.

Keywords: Distance teaching, teachware, Internet, multimedia

This research have been done thanks to a Eurostat project (AC/15/97/OSCE)

1 Distance Teaching and Video

Distance teaching has been developed in many countries such as France by the CNED, UK with the Open University, Australia with the " TV Open learning project ", Japan with the University of the Air, (cf. Hayashi) among many others. All these institutions propose courses in statistics, using one or several technologies: paper and surface mail, audiotapes, videotapes, computer software. What is typical of these institutions is not only the technologies, but the teaching organisation with contacts with a tutor, meetings, phone conferences.
Videotapes are commonly used in Australia, Japan. In the english speaking world, the series " Against all Odds " produced by The Annenberg Corporation for Public Broadcasting has been widely used: directed by D.Moore, (cf. Moore) it consists into 26 lectures of 30 minutes plus 70 documentary segments about one topic of application..
The NSF has also produced in 1992 the series " Statistics: decisions through data " for the secondary school. The " ASA short courses-tutorial videotapes " are of a much higher level with eg a 12 hours course on multivariate analysis.
In France, we may notice the video course by D.Schwartz (cf. Schwartz) for introducing variability to biologists.

2 Teachwares

Software for Computer Assisted Training in Statistics may be roughly divided into three categories. Even if some are available for demonstration purposes on the Web, the products presented here are not internet based in the sense of part 3.

2.1 Software based on spreadsheets

Spreadsheets like Lotus or Excel present many advantages: they are widespread, and one may perform a great variety of statistical analysis on one's own data. Usually teachware based on spreadsheets are very cheap: 10$ for XLStatistics of R.Carr from Deakin University (http://www.man.deakin.edu.au/rodneyc/) one can find topics like Cumulative probability plots, Scatterplots with axes swapped, regression for straight lines with a possible break-point, Fitting a smooth curve to data (Moving average, Running median, Means of grouped data, Locally weighted linear regression, etc.
DISCUS is another package developped at Coventry University (http://www.mis.coventry.ac.uk/research/discus/discus_home.html), it is user-friendly and in one diskette it covers a whole course of introductory statistics for £23.5. with Descriptive Statistics}, Probability, Binomial Distributions, the

Poisson Distribution, Continuous Distributions, Sampling, Regression, Hypothesis Testing,

2.2 Products based upon professional statistical software

Let us just mention that several statististical software have a learning module which is a compromise between a on-line help and a statistical course see SAS-Assist or SURVO (http://www.helsinki.fi/survo/survoe.html).

2.3 Specific software for teaching purposes

This is an increasing category where software is written with special development tools for educational purpose. Let us mention a few products:
STEPS (Statistical Education through Problem Solving) produced by CTI Statistics (http://www.stats.gla.ac.uk/steps/home.html), is mainly concerned with applications in social sciences and biology. It is based on Toolbook, Visual Basic and XlispStat.
QUERCUS is another product developed by a consortium of Scottish Universities (http://www.stams.strath.ac.uk/external/CAL/quercus/) written in Authorware Pro. Quercus is a suite of interactive tutorial software designed to tutor bioscience students in the basic techniques of statistical analysis statistics. Supplementary modules (ExerciseBooks),compliment the WorkBooks by providing exercises to reinforce learning and self-assessment tests.
CNAM Media Statistique (cf. Cnam-Media) is a french product written in Toolbook, developped since 1990. It covers a course in introductory statistics. It is an interactive course where each module is structured into a presentation, a development, a summary and a set of exercises. Each screen is splitted into two parts: text on the left part, graphics and animation in the right part. Each module corresponds to a self learning of several hours. Here are some typical screens:
P.Dassonville (cf. Dassonville) has developped a real multimedia software, including voice, animation, etc. PRAME, with the concept of a virtual campus. This product, one of the most advanced, is still a prototype dealing with only one method (principal components) and available in the network of french business schools.
There are many other packages like Stateasy from A.Di Ciaccio and S.Borra, E-STAT from Statistics Canada. Almost all the packages presented here need a standard PC configuration with Windows3.1.

3 Internet based products

One of the most ambitious product is the UCLA electronic statistics textbook (Ucla-Est) called ″Statistics. The study of stability in variation″ developed by J. de Leeuw (cf. J. de Leeuw) and his team (http://www.stat.ucla.edu/textbook/). The goal of the authors is to write a textbook which is freely available to everyone on the internet, independent of the level of the student, interactive, using graphics and demos and which covers most of the statistical theory as traditionally taught. The contents of the book is traditional with an introduction, the analysis of a single variable, the analysis of a pair of variables and the analysis of several variables. The book is already under construction but nevertheless it is useful and user-friendly. It provides at any time a statistics glossary, a statistics toolbox and Xlisp-Stat demos. Besides, many very appealing applets are available, written by B. Narasimhan (Stanford), Webster West (SC), David Lane (Rice), Tony Rossini (SC), and others.

On the other hand, David Lane (http://www.ruf.rice.edu/~lane/hyperstat/) wrote HyperStat in HyperCard a couple of years ago. Hyperstat on-line is an introductory-level hypertext statistics book.

Let‘s mention also the Chance database which contains materials designed to help teach a Chance course or a more standard introductory probability or statistics course. (http://www.dartmouth.edu/~chance/). The Chance course is a case study quantitative literacy course developed cooperatively by Middlebury, Grinnell, Spelman, University of California San Diego, University of Minnesota, and Dartmouth. The aim of Chance is to make students more informed, and critical, readers of current news that uses probability and statistics as reported in daily newspapers such as "The New York Times" and current journals and magazines such as "Chance", "Science", "Nature", and the "New England Journal of Medicine".

The goal of the project called ″Virtual Laboratories in Probability and statistics″ is to provide interactive, web-based modules for students and teachers of probability and statistics (http://www.math.uah.edu/~stat/). It covers special models such as geometric models, Bernoulli trials and so on, basic probability and basic statistics.

The Claremont Colleges' "Web Interface for Statistics Education" seeks to expand teaching resources offered through Introductory Statistics courses, especially in the social sciences. (http://www.grad.cgs.edu/wise/) . This project aims to develop an on-line teaching tool to take advantage of the unique hypertextual and presentational benefits of the World Wide Web (WWW). This teaching tool's primary application is as a supplement to traditional teaching materials, addressing specific topics that instructors have difficulty in presenting using traditional classroom technologies. The tool serves to promote self-paced learning and to provide a means for advanced students to review concepts.

XploRe is an highly interactive computing environment which combines the application of statistical procedures as well as the development of algorithms for

statistical exploration of data (http://www.xplore-stat.de/). Among the main features of XploRe, are fast programming language, graphical user interface, network facilities, and self-driven help system. It provides also tutorials in statistics , and in programming. (cf. M. Muller)
The above list is not exhaustive and products are changing very fast. Some updated lists of products are available, for instance the following one http://noppa5.pc.helsinki.fi/links.html which is supported by J. Puranen.

4- Conclusion

There are a lot of available products for self-learning in statistics but only a few of them are interactive and cover a large statistical area : CNAM-Media, Steps, Ucla-Est. In France, PRAME is a real multimedia educational sofware in data analysis but to date, it exists only as a prototype.
Most of the products can be downloaded from the developer site and used in an independant way by the client. However, some products, allow an interaction between the developer and the user (UCLA-est for instance). Then, if one starts a program from the HTML page, then in some cases it is executed on the server (i.e. the machine of the product developer), and in some cases on the client (i.e. his/her machine). Both options have advantages and disadvantages. If the program is executed on the server, then one does not have to assume anything about the client, and one can set things up so that they work correctly. But lots of users will cause a lot of load on the developer machine. And since the results have to be sent over the network to the client, the response may be a bit slow. On the other hand, if the program is executed on the client machine, it has to be downloaded first (in the case of a Java applet), and the developer is unable to help the client if his/her setup is not correct. Similarly, if the client browser cannot support Java, JavaScript, client-pull, server-push, and so on, there is no way we can help.
Another risk of multimedia as mentioned in ″Multimedia and new educational environments″ (cf. M G. Ottaviani) is that the student can get lost with all the sequential links. Therefore the design of interface must be convenient and take into account the problem of dealing with hypertext and/or hypermedia.
At the end, let us mention several points. First , we don‘t know any study evaluating multimedia statistical learning packages . It could be valuable to define some criteria (minimum desirable such as quality, efficiency) that such a product must reach. Besides, it could be interesting to keep a trace of the student ″journey ″ through the hypermedia course. When they offer an assessment, most of the products are self-evaluating. The insertion of an evaluation device which can be studied by the teacher increases heavily the complexity of the product (cf. Morin) but it is essential to improve the product. To conclude, the best hypermedia technology will never replaced human teachers (cf. Velleman and Moore) and they must be designed supplementary to each other.

References

Cnam-Media Statistique, Jeriko, 13 rue Vernier, 75017 Paris

P. Dassonville Presentation of a CD-ROM for learning correlation and PCA PRAME , 1997, ISI 51st meeting Istanbul.

J. de Leeuw The UCLA Statistics Textbooks and Modules 1997, ISI 51st meeting Istanbul pp 55-59

A DiCiaccio : Teaching basic statistics following a multimedia approach 1996, in Multimedia and new educational environments Ed MG Ottaviani 44-54

C. Hayashi : Teaching Statistics and Data analysis at the University of the Air 1994, Proc. of ICOTS-4, Marrakech, 170-177.

D.S. Moore: The place of video in new styles of teaching and learning statistics 1993 The American Statistician vol 47, pp 172-176.

M. Muller Teaching Statistics with XploRe 1998 Newsletter Maths &Stats

A. Morin Logiciel: Mathématiques par thèmes, Module de statistique. 1994, CNED Rennes.

M.G.Ottaviani,(ed), Multimedia and new educational environments, Dipartimento di Statistica, Universita di Roma La Sapienza,1996, 80 pages.

D. Schwartz : Les statistiques dans les sciences du vivant,1991, INSERM,

M Talbot : SMART : Introducing Specialist Statistical Techniques via the Web 1998 Newsletter Maths &Statistics.

Velleman, D.S. Moore : Multimedia for teaching statistics : promises and pitfalls, 1996 The American Statistician Vol 50, 3 , 217-225.